```
GGCTGAGCTGCTGCCGCCGCCGGGCGGACGGGCGGACGCGCGGAGCTGGGGGCGGCGCGGCGGGGCCGGCGGGGCGCGG

                      M   R   R   K   E   K   R   L   L   Q   A   V   A   L   V     15
CGGGGCTGACCGGCCCCG ATG AGG CGG AAG GAG AAG CGG CTC CTG CAG GCG GTG GCG CTG GTG    45

 L   A   A   L   V   L   L   P   N   V   G   L   W   A   L   Y   R   E   R   Q     35
CTG GCG GCC CTG GTC CTC CTG CCC AAC GTG GGG CTT TGG GCG CTG TAC CGC GAG CGG CAG   105

 P   D   G   T   P   G   G   S   G   A   A   V   A   P   A   A   G   Q   G   S     55
CCC GAC GGC ACC CCT GGG GGA TCG GGG GCG GCG GTG GCG CCG GCG GCG GGA CAG GGC TCA   165

 H   S   R   Q   K   K   T   F   F   L   G   D   G   Q   K   L   K   D   W   H     75
CAC AGT CGA CAA AAG AAA ACG TTT TTC TTG GGA GAT GGG CAG AAG CTG AAG GAC TGG CAT   225

 D   K   E   A   I   R   R   D   A   Q   R   V   G   N   G   E   Q   G   R   P     95
GAC AAG GAG GCC ATC CGG AGG GAC GCT CAG CGC GTA GGA AAT GGA GAA CAA GGA AGA CCT   285

 Y   P   M   T   D   A   E   R   V   D   Q   A   Y   R   E   N   G   F   N   I    115
TAC CCC ATG ACC GAT GCT GAG AGA GTG GAT CAG GCA TAC CGA GAA AAT GGA TTT AAC ATC   345

 Y   V   S   D   K   I   S   L   N   R   S   L   P   D   I   R   H   P   N   C    135
TAC GTC AGT GAT AAA ATC TCC TTG AAT CGC TCT CTC CCA GAT ATC CGG CAC CCA AAC TGC   405

 N   S   K   R   Y   L   E   T   L   P   N   T   S   I   I   I   P   F   H   N    155
AAC AGC AAG CGC TAC CTG GAG ACA CTT CCC AAC ACA AGC ATC ATC ATC CCC TTC CAC AAC   465
```

Fig. 1A

```
 E   G   W   S   S   L   L   R   T   V   H   S   V   L   N   R   S   P   P   E   175
GAG GGC TGG TCC TCC CTC CTC CGC ACC GTC CAC AGT GTG CTC AAT CGC TCG CCT CCA GAG   525

 L   V   A   E   I   V   L   V   D   D   F   S   D   R   E   H   L   K   K   P   195
CTG GTC GCC GAG ATT GTA CTG GTC GAC GAC TTC AGT GAT CGA GAG CAC CTG AAG AAG CCT   585

 L   E   D   Y   M   A   L   F   P   S   V   R   I   L   R   T   K   K   R   E   215
CTT GAA GAC TAC ATG GCC CTT TTC CCC AGT GTG AGG ATT CTT CGA ACC AAG AAA CGG GAA   645

 G   L   I   R   T   R   M   L   G   A   S   V   A   T   G   D   V   I   T   F   235
GGG CTG ATA AGG ACC CGA ATG CTG GGG GCC TCA GTG GCA ACT GGG GAT GTC ATC ACA TTC   705

 L   D   S   H   C   E   A   N   V   N   W   L   P   P   L   L   D   R   I   A   255
TTG GAT TCA CAC TGT GAA GCC AAT GTC AAC TGG CTT CCC CCC TTG CTT GAC CGC ATT GCT   765

 R   N   R   K   T   I   V   C   P   M   I   D   V   I   D   H   D   D   F   R   275
CGG AAC CGC AAG ACC ATT GTG TGC CCG ATG ATT GAT GTA ATT GAC CAT GAC GAC TTT CGG   825

 Y   E   T   Q   A   G   D   A   M   R   G   A   F   D   W   E   M   Y   Y   K   295
TAC GAG ACA CAG GCA GGG GAT GCC ATG CGG GGA GCC TTT GAC TGG GAG ATG TAC TAC AAG   885

 R   I   P   I   P   P   E   L   Q   K   A   D   P   S   D   P   F   E   S   P   315
CGG ATC CCG ATC CCT CCA GAA CTG CAG AAA GCT GAC CCC AGC GAC CCA TTT GAG TCT CCC   945

 V   M   A   G   G   L   F   A   V   D   R   K   W   F   W   E   L   G   G   Y   335
GTG ATG GCC GGT GGA CTG TTC GCC GTG GAT CGG AAG TGG TTC TGG GAA CTC GGC GGG TAT 1005
```

Fig. 1B

PATENT DRAFT

```
 D   P   G   L   E   I   W   G   G   E   Q   Y   E   I   S   F   K   V   W   M   355
GAC CCA GGC TTG GAG ATC TGG GGA GGG GAG CAG TAT GAA ATC TCC TTC AAG GTG TGG ATG 1065

 C   G   G   R   M   E   D   I   P   C   S   R   V   G   H   I   Y   R   K   Y   375
TGT GGG GGC CGC ATG GAG GAC ATC CCC TGC TCC AGG GTG GGC CAT ATC TAC AGG AAG TAT 1125

 V   P   Y   K   V   P   A   G   V   S   L   A   R   N   L   K   R   V   A   E   395
GTG CCC TAC AAG GTC CCG GCC GGA GTC AGC CTG GCC CGG AAC CTT AAG CGG GTG GCC GAA 1185

 V   W   M   D   E   Y   A   E   Y   I   Y   Q   R   R   P   E   Y   R   H   L   415
GTG TGG ATG GAT GAG TAC GCA GAG TAC ATT TAC CAG CGC CGG CCT GAA TAC CGC CAC CTC 1245

 S   A   G   D   V   A   V   Q   K   K   L   R   S   S   L   N   C   K   S   F   435
TCC GCT GGG GAT GTC GCA GTC CAG AAA AAG CTC CGC AGC TCC CTT AAC TGC AAG AGT TTC 1305

 K   W   F   M   T   K   I   A   W   D   L   P   K   F   Y   P   P   V   E   P   455
AAG TGG TTT ATG ACG AAG ATA GCC TGG GAC CTG CCC AAA TTC TAC CCA CCC GTG GAG CCC 1365

 P   A   A   A   W   G   E   I   R   N   V   G   T   G   L   C   A   D   T   K   475
CCG GCT GCA GCT TGG GGG GAG ATC CGA AAT GTG GGC ACA GGG CTG TGT GCA GAC ACA AAG 1425

 H   G   A   L   G   S   P   L   R   L   E   G   C   V   R   G   R   G   E   A   495
CAC GGG GCC TTG GGC TCC CCA CTA AGG CTA GAG GGC TGC GTC CGA GGC CGT GGG GAG GCT 1485

 A   W   N   N   M   Q   V   F   T   F   T   W   R   E   D   I   R   P   G   D   515
GCC TGG AAC AAC ATG CAG GTA TTC ACC TTC ACC TGG AGA GAG GAC ATC CGG CCT GGA GAC 1545
```

Fig. 1C

Title: "47169 and 33935, Novel Human Glycosyl Transferases and Uses Thereof"
Inventors: Rachel A. Meyers et al.
U.S. Patent Appl. No.: **Not Yet Assigned**
Express Mail # EL916936575US Attorney Docket No. 10147-56U1 Cust # 570

```
 P   Q   H   T   K   K   F   C   F   D   A   I   S   H   T   S   P   V   T   L   535
CCC CAG CAC ACC AAG AAG TTC TGC TTT GAT GCC ATT TCC CAC ACC AGC CCT GTC ACG CTG 1605

 Y   D   C   H   S   M   K   G   N   Q   L   W   K   Y   R   K   D   K   T   L   555
TAC GAC TGC CAC AGC ATG AAG GGC AAC CAG CTG TGG AAA TAC CGC AAA GAC AAG ACC CTG 1665

 Y   H   P   V   S   G   S   C   M   D   C   S   E   S   D   H   R   I   F   M   575
TAC CAC CCT GTC AGT GGC AGC TGC ATG GAC TGC AGT GAA AGT GAC CAT AGG ATC TTC ATG 1725

 N   T   C   N   P   S   S   L   T   Q   Q   W   L   F   E   H   T   N   S   T   595
AAC ACC TGC AAC CCA TCC TCT CTC ACC CAG CAG TGG CTG TTT GAA CAC ACC AAC TCA ACA 1785

 V   L   E   K   F   N   R   N   *                                               603
GTC TTG GAA AAA TTC AAT AGG AAC TGA                                             1812

GCCCTCATGTCCCCTTGGCAGGCCCCCCAGGGTCTGGCACTCACTGCAGACTTCCTCTTTCAAGGGAGGCAGGGCCCCT
GTGGGCACTAGGTGTAAAAGGTGCTGGCCAAATGGTTCAGGGTGAAGAGGGCTCTTGATTCAGGGGCTGGGGTCTGCCT
GGTCCTTGAGCCCCTGAGTTGTGGGGGTAGGGTGAAGAGCATATCCCACAAGAGGCCCCACAGGGAGCAGAGACTGCTT
TAATCCCTGCTGACATCACGGAAAAGCAACAGAGCCTTTTCAACTTTGTCACTATGTCCCCTTGAACATTATGTGGGAG
AACACCAAGGTAGCCTAGGCCACCCAAAAGTGAGTCCTGCGAGGTTGCCCAGCCCTCAGATGGCTCTCCTACATGATGG
TGCTTTAGAAACAAAGGTAAAATTTGCCTGTTTGGGGCAGCTTTTAGTATCGATGCCACTCATCTGCAGCAGAAGAGAA
AGAAGTCCTCTTGGGGCTTTTTAGTTTCTGCCGTCCTGGGGGGAACATTGCAGTTACTGCACAGCTTCTGTTCTCTGTC
ACAACCCCAGGTGATTTGGTCCGGTCAAAGGCCATACTTGGGGCCCTAAGAGTGTTCAGTATTGAATGCTGATCAGCTG
CCAGGTGAGGAGTCAGAAGAGGGAGCCCCCCTAGACATTTCTTTGCAGCTATGGACATGCGGGATATCTCCCCCTGCTC
TCTGGGTATTTGAAATGTCAATTTTAGCACTCTCCAGGCACAAGGACAGCCCAGCACCAGCTTTACAGGGCAGTGTTTC
```

Fig. 1D

AGATGGCCCTGAGCCCACGGAAAAGGCCAGGTAGACCTCCAAACTAGAAATGCTGGCTGATTTGCCCTGATCCATGCTT
CCATTTCCCTGTCTCTCTTCCCCAGGCAATTACTGGCCTCAAAAGAGGAACAGAGGTGCTGCGAGGTGCTCACCTCACA
GAGTCTGGAGGCCTCCAGGATCAACTGTGGGCAAAGTGCCTGCCTCTGACCTCATCATGGTTCTAGTTCTCATACAGAA
CTCCAGAATTTTTAAAGAACTCTATAATTGGATTGCAAACTAGGATGCTACATAGGATTCTGGTATTCCACATCCAATA
TGGATTTCTAGAATGCTGTGATTAAAGGAGCCAGCCAGGTGTAATACAGTCAAGGCAGCCCCCAGCCTAGAGACAATCT
GTGAAATCCAAAGTTGGTGGTGTTGGGAAAGCAGGGGGACATGTGTCCCTCAGCTCAGCAGAGGCTGTGGTACAACATG
GTCCTTGGTGAAGACCTGCACCCCTGGAACCTCCCACCATCATCACAACTGTAGTCTCATTTGCAGTGGAGAAAAGAAC
CCGACGTCCCACAGCCAGATATACACCCAGCTCCATGCCAGCCCTTCATGTTTACCTTTTGCTTTGTTAATTACATGTC
AGACTCCTAGAGGGCCTCCAGACTAATAGGAAGCATTTCTGTAACCAACCTGCCACCCACTGATTCAGAAATGGAAATC
ACATTCCACAATCTATGGCTTCCACCAGCTAGCCCAGGAAATACTTGAAATCAGCATTCCAATTAGTGTTGAGTCTCTT
GATTGTGTCATTTACCAATTAAATAACTGAGACCTAAGTCTGGGAACAGAGCCACGAATCTGCCTTTGAGATGCTGGCA
GATCTCAAGGCCATCAATTATTGGGGGAGGGAGGGACAAACACTCCCAATCATCCACCAGTCAGACTGAATGTGTAGCT
GGCGAGGAATTACTTCCACTTCTGGCCCAGCACAAGCCCTGCTTTGGCCACCTGTCTGCAAGAGAGGCGGCCCCTGTGC
TTGCAACGCTTACGTGTTGATCCCAGTGTCCTTTTCCAAATGAGTGCTGTAGCTTTAGAAGTGGCCCTCTATAGAAAGA
AGTCAAAAGATGAGGCCCCTTCTAGAATCTAGGATAACAAGAGTGTTGACAGTTTGAGGAGTCGAATTGAGATTCATCA
TCAAAGAGCAATGCAGCGTCGTTAAAATAAAAACTGTGCCTTTTAAAAAGAAAAATGCAAATATAGAGCAAATCCCTAA
ACTTGAAAAAAAAAAAAAAAA

Fig. 1E

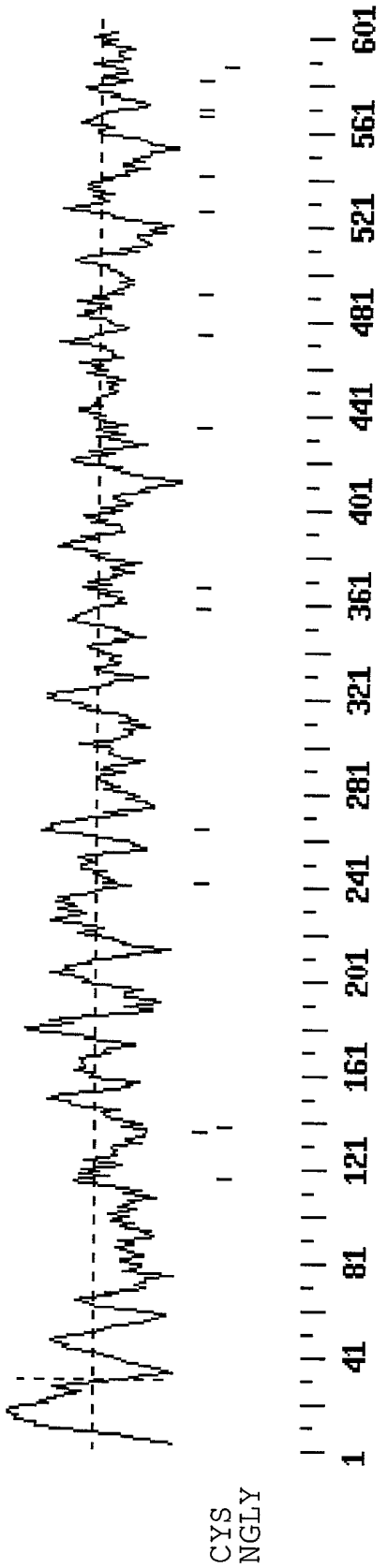
CYS
NGLY
1
41
81
121
161
201
241
281
321
361
401
441
481
521
561
601

Fig. 2

```
                M   A   A   G   E   R   S   W   C   L   C   K   L   L   R   F   F     17
TCGGCGGAAG ATG GCG GCC GGC GAA AGG AGC TGG TGC CTG TGC AAG TTG TTG AGG TTT TTT    51

 Y   S   L   F   F   P   G   L   I   V   C   G   T   L   C   V   C   L   V   I    37
TAT TCA TTA TTC TTC CCT GGG CTC ATT GTA TGT GGA ACT TTA TGT GTG TGT TTG GTC ATT  111

 V   L   W   G   I   R   L   L   L   Q   R   K   K   K   L   V   S   T   S   K    57
GTC CTT TGG GGA ATC AGA CTG CTG CTA CAG AGA AAG AAA AAA TTA GTG TCA ACT AGC AAA  171

 N   G   K   N   Q   M   V   I   A   F   F   H   P   Y   C   N   A   G   G   G    77
AAT GGG AAA AAT CAA ATG GTG ATT GCA TTT TTT CAT CCA TAC TGC AAT GCT GGT GGA GGA  231

 G   E   R   V   L   W   C   A   L   R   A   L   Q   K   K   Y   P   E   A   V    97
GGA GAA AGA GTT TTA TGG TGT GCT TTA AGA GCC CTG CAG AAA AAG TAT CCT GAA GCA GTT  291

 Y   V   V   Y   T   G   D   V   N   V   N   G   Q   Q   I   L   E   G   A   F   117
TAT GTT GTT TAT ACC GGC GAT GTT AAT GTC AAC GGT CAA CAG ATA CTA GAA GGT GCT TTC  351

 R   R   F   N   I   R   L   I   H   P   V   Q   F   V   F   L   R   K   R   Y   137
AGA AGA TTT AAC ATC AGA TTA ATT CAC CCA GTG CAG TTT GTT TTT TTA AGG AAA CGC TAT  411

 L   V   E   D   S   L   Y   P   H   F   T   L   L   G   Q   S   L   G   S   I   157
CTT GTG GAA GAT TCA CTG TAT CCT CAC TTC ACA CTG CTG GGC CAA AGT CTA GGA TCC ATT  471

 F   L   G   W   E   A   L   M   Q   C   V   P   D   V   Y   I   D   S   M   G   177
TTT CTT GGC TGG GAA GCT CTA ATG CAG TGT GTT CCT GAT GTT TAC ATT GAT TCA ATG GGA  531
```

Fig. 3A

Y	A	F	T	L	P	L	F	K	Y	I	G	G	C	Q	V	G	S	Y	V	197
TAC	GCT	TTT	ACG	CTT	CCT	CTG	TTT	AAG	TAT	ATA	GGG	GGT	TGC	CAA	GTT	GGA	AGC	TAT	GTT	591
H	Y	P	T	I	S	T	D	M	L	S	V	V	K	N	Q	N	I	G	F	217
CAT	TAT	CCT	ACT	ATC	AGC	ACC	GAC	ATG	CTC	TCT	GTA	GTG	AAG	AAT	CAA	AAT	ATT	GGA	TTT	651
N	N	A	A	F	I	T	R	N	P	F	L	S	K	V	K	L	I	Y	Y	237
AAT	AAT	GCA	GCC	TTC	ATT	ACC	AGG	AAT	CCT	TTT	CTC	AGC	AAA	GTA	AAG	CTC	ATC	TAC	TAC	711
Y	L	F	A	F	I	Y	G	L	V	G	S	C	S	D	V	V	M	V	N	257
TAT	TTA	TTT	GCT	TTT	ATT	TAT	GGA	CTT	GTT	GGT	TCT	TGC	AGT	GAT	GTA	GTC	ATG	GTC	AAT	771
S	S	W	T	L	N	H	I	L	S	L	W	K	V	G	N	C	T	N	I	277
TCT	TCT	TGG	ACA	CTA	AAC	CAT	ATT	CTC	TCA	CTA	TGG	AAA	GTT	GGG	AAT	TGC	ACT	AAC	ATT	831
V	Y	P	P	C	D	V	Q	T	F	L	D	I	P	L	H	E	K	K	M	297
GTT	TAT	CCA	CCT	TGT	GAT	GTG	CAG	ACA	TTT	CTG	GAC	ATT	CCC	TTA	CAT	GAG	AAA	AAG	ATG	891
T	P	G	H	L	L	V	S	V	G	Q	F	R	P	E	K	N	H	P	L	317
ACC	CCA	GGA	CAT	TTG	CTG	GTT	TCT	GTT	GGC	CAG	TTT	AGG	CCG	GAA	AAG	AAT	CAT	CCA	TTG	951
Q	I	R	A	F	A	K	L	L	N	K	K	M	V	E	S	P	P	S	L	337
CAG	ATC	AGA	GCC	TTT	GCT	AAA	TTG	CTG	AAT	AAG	AAG	ATG	GTT	GAG	TCA	CCT	CCT	TCG	CTT	1011
K	L	V	L	I	G	G	C	R	N	K	D	D	E	L	R	V	N	Q	L	357
AAA	CTT	GTC	CTC	ATT	GGA	GGT	TGT	CGT	AAC	AAA	GAT	GAT	GAA	CTT	AGG	GTA	AAC	CAA	CTG	1071

Fig. 3B

```
 R   R   L   S   E   D   L   G   V   Q   E   Y   V   E   F   K   I   N   I   P   377
AGA AGG CTG TCT GAG GAT TTA GGA GTT CAA GAA TAT GTG GAA TTT AAA ATA AAC ATT CCA 1131

 F   D   E   L   K   N   Y   L   S   E   A   T   I   G   L   H   T   M   W   N   397
TTT GAT GAA TTA AAG AAT TAT TTG TCT GAA GCA ACA ATT GGT CTG CAT ACC ATG TGG AAC 1191

 E   H   F   G   I   G   V   V   E   C   M   A   A   G   T   I   I   L   A   H   417
GAG CAT TTT GGG ATT GGA GTT GTG GAG TGT ATG GCA GCT GGC ACA ATT ATC CTT GCA CAC 1251

 N   S   G   G   P   K   L   D   I   V   V   P   H   E   G   D   I   T   G   F   437
AAT TCG GGG GGC CCA AAG CTT GAC ATT GTG GTT CCT CAC GAA GGA GAT ATA ACT GGC TTT 1311

 L   A   E   S   E   E   D   Y   A   E   T   I   A   H   I   L   S   M   S   A   457
CTG GCT GAG AGT GAA GAA GAC TAT GCT GAA ACT ATC GCT CAC ATT CTT TCC ATG TCT GCA 1371

 E   K   R   L   Q   I   R   K   S   A   R   A   S   V   S   R   F   S   D   Q   477
GAA AAG AGA CTC CAA ATC AGA AAA AGT GCT CGT GCA TCT GTA AGC AGA TTC TCT GAT CAG 1431

 E   F   E   V   T   F   L   S   S   V   E   K   L   F   K   *                   492
GAA TTT GAA GTG ACA TTC CTA TCA TCT GTG GAA AAG TTA TTT AAG TAA                   1479

TGCCATATCTGTAAAATTAAAGATATTTTATATAAACTGGTTAAACACCTTCATATGTAAATATTTTTCTAAATTCAAT
CTCATTTGTCAAATCATTTTACTTTAGAAAACAGACAAAATTTCCTTTTAGAATAAAAGGAAGTGTTGAAAAGAAAATG
GATGACTAGCCTTCGGCTTCCATTCTTGGTATACATGAGAGAGGCTGGCTGCTGAGATGAATGTGAACCAGGTTGCAGA
GAATCTGGCTTTGAGCCACCAGGAAGAACTAGTGGATTTGCCAAAAAACTACCCCTTGAGTGAAAATGAAGATGAGGGG
```

Fig. 3C

GACAGTGATGGAGAGAGAAAGCATCAAAAGCTTCTGGAAGCAATCATTTCCCTTGATGGAAAGAATAGGCGGAAATTGG
CTGAGAGGTCTGAGGCTAGTCTGAAAGTGTCAGAGTTCAGTGTCAGTTCTGAAGGATCAGGAGAAAAGCTGGGCCTTGC
AGATCTGCTTGAGCCCGTTAAAACTTCATCTTCTTTGGCCACTGTAAAAAAGCAACTGAATAGAGTCAAATCAAAGAAG
GTGGTGGAGTTACCTCTTAACAAAGAAAAAATTGAACAGATCCACAGAGAAGTAGCATTCAGTAAAACCTCACAGGTCC
TCTCCAAATGGGACCCTATCATCCTGAAGAACCAGCAGGCAGAGCAGCTGGTTTTTCCCCTGGGGAAGGAGCAGCCAGC
CATTGCTCCCATTGAACATGCGCTCAGTGGCTGGAAGGCAAGAACTCCCCTGGAGCAGGAAATTTTTAACCTCCTCCAT
AAGAACAAGCAGCCAGTGACAGATCCTTTACTGACTCCCATGGAAAAGGCCTCTCTCCAAGCCATGAGCCTGGAAGAGG
CAAAGATGCACCGAGCAGAGCTTCAGAGGGCTCGGGCTCTGCAGTCCTACTATGAGGCCAAGGCTCGAAAAGAGAAGAA
AATCAAAAGTAAAAAGTATCACAAAGTCGTGAAGAAAGGAAAGGCCAAGAAAGCCTTAAAAGAGTTTGAGCAGCTACAG
AAGGTTAATCCAACTGTGGCACTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Fig. 3D

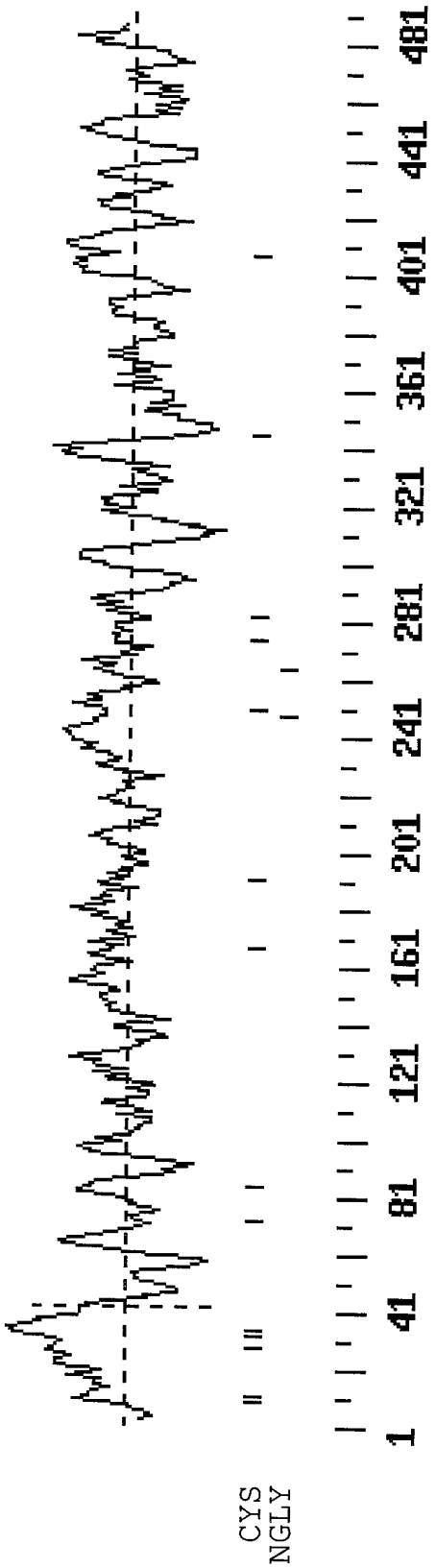

Fig. 4

```
MUS_GT1 MRKFAYCKVVLATSL VWVLLDMFLLLYFSE CNKCEEKQERGLPAG ---------------
RAT_GT1 MRKFAYCKVVLATSL VWVLLDMFLLLYFSE CNKCEEKKERGLPAG ---------------
HUM_GT1 MRKFAYCKVVLATSL IWVLLDMFLLLYFSE CNKCDEKKERGLPAG ---------------
COW_GT1 MRKFAYCKVVLATSL IWVLLDMFLLLYFSE CNKCDEKKERGLPAG ---------------
PIG_GT1 MRKFAYCKVVLATSL IWVLLDMFLLLYFSE CNKCDEKKERGLPAG ---------------
47169   ---MRRKEKRLLQAV ALVLAALVLLPNVGL WALYRERQPDGTPGG SGAAVAPAAGQGSHS
                  :...: :.::: :.::  .       :::  : :::

MUS_GT1 ---------------- --DVLELVQKPHEGP GEMGKPVVIPKEDQE KMKEMFKINQFNLMA
RAT_GT1 ---------------- --DVLELVQKPHEGP GEMGKPVVIPKEDQE KMKEMFKINQFNLMA
HUM_GT1 ---------------- --DVLEPVQKPHEGP GEMGKPVVIPKEDQE KMKEMFKINQFNLMA
COW_GT1 ---------------- --DVLEPVQKPHEGP GEMGKPVVIPKEDQE KMKEMFKINQFNLMA
PIG_GT1 ---------------- --DVLEPVQKPHEGP GEMGKPVVIPKEDQD KMKEMFKINQFNLMA
47169   RQKKTFFLGDGQKLK DWHDKEAIRRDAQRV GNGEQGRPYPMTDAE RVDQAYRENGFNIYV
                       :.:::      : : :.  .:  : : ::..::: : ::: :

MUS_GT1 SEMIALNRSLPDVRL EGCKTKVYPDNLPTT SVVIVFHNEAWSTLL RTVHSVINRSPRHMI
RAT_GT1 SEMIAFNRSLPDVRL EGCKTKVYPDSLPTT SVVIVFHNEAWSTLL RTVHSVINRSPRHMI
HUM_GT1 SEMIALNRSLPDVRL EGCKTKVYPDNLPTT SVVIVFHNEAWSTLL RTVHSVINRSPRHMI
COW_GT1 SEMIALNRSLPDVRL EGCKTKVYPDNLPTT SVVIVFHNEAWSTLL RTVHSVINRSPRHML
PIG_GT1 SEMIALNRSLPDVRL EGCKTKVYPDNLPTT SVVIVFHNEAWSTLL RTVHSVINRSPRHML
47169   SDKISLNRSLPDIRH PNCNSKRYLETLPNT SIIIPFHNEGWSSLL RTVHSVLNRSPPELV
        :: :.:::::::::    :::: :::.::.: ::::.:::::::::: ::::::::::::.::
```

Fig. 5A

```
MUS_GT1 EEIVLVDDASERDFL KRPLESYVKKLKVPV HVIRMEQRSGLIRAR LKGAAVSRGQVITFL
RAT_GT1 EEIVLVDDASERDFL KRPLESYVKKLKVPV HVIRMEQRSGLIRAR LKGAAVSKGQVITFL
HUM_GT1 EEIVLVDDASERDFL KRPLESYVKKLKVPV HVIRMEQRSGLIRAR LKGAAVSKGQVITFL
COW_GT1 EEIVLVDDASERDFL KRPLESYVKKLKVPV HVIRMEQRSGLIRAR LKGAAVSKGQVITFL
PIG_GT1 EEIVLVDDASERDFL KRPLESYVKKLKVPV HVIRMEQRSGLIRAR LKGAAVSKGQVITFL
47169   AEIVLVDDFSDREHL KKPLEDYMALFPS-V RILRTKKREGLIRTR MLGASVATGDVITFL
          ::::::: :::::: ::::::.::   .   : :::: .::.:::: : : :: :  :.::::::

MUS_GT1 DAHCECTAGWLEPLL ARIKHDRRTVVCPII DVISDDTFEYMAGS- DMTYGGFNWKLNFRW
RAT_GT1 DAHCECTVGWLEPLL ARIKHDRRTVVCPII DVISDDTFEYMAGS- DMTYGGFNWKLNFRW
HUM_GT1 DAHCECTVGWLEPLL ARIKHDRRTVVCPII DVISDDTFEYMAGS- DMTYGGFNWKLNFRW
COW_GT1 DAHCECTVGWLEPLL ARIKHDRKTVVCPII DVISDDTFEYMAGS- DMTYGGFNWKLNFRW
PIG_GT1 DAHCECTVGWLEPLL ARIKHDRKTVVCPII DVISDDTFEYMAGS- DMTYGGFNWKLNFRW
47169   DSHCEANVNWLPPLL DRIARNRKTIVCPMI DVIDHDDFRYETQAG DAMRGAFDWEMYYKR
        : :::..: :: :::  :: :.:::::::::: :::..:.:.:       ::  :::.:.: :::

MUS_GT1 YPVPQREMDRRKGDR TLPVRTPTMAGGLFS IDRDYFQEIGTYDAG MDIWGGENLEISFRI
RAT_GT1 YPVPQREMDRRKGDR TLPVRTPTMAGGLFS IDRDYFQEIGTYDAG MDIWGGENLEISFRI
HUM_GT1 YPVPQREMDRRKGDR TLPVRTPTMAGGLFS IDRDYFQEIGTYDAG MDIWGGENLEISFRI
COW_GT1 YPVPQREMDRRKGDR TLPVRTPTMAGGLFS IDRDYFQEIGTYDAG MDIWGGENLEISFRI
PIG_GT1 YPVPQREMDRRKGDR TLPVRTPTMAGGLFS IDRDYFQEIGTYDAG MDIWGGENLEISFRI
47169   IPIPP---ELQKADP SDPFESPVMAGGLFA VDRKWFWELGGYDPG LEIWGGEQYEISFKV
        .:::    : ::::  : :..:: ::::::  :::.:::::: ::.: :::::::::.::::::
```

Fig. 5B

```
MUS_GT1 WQCGGTLEIVTCSHV GHVFRKATPYTFPGG TGQIINKNNRRLAEV WMDEFKNFFYIISPG
RAT_GT1 WQCGGTLEIVTCSHV GHVFRKATPYTFPGG TGQIINKNNRRLAEV WMDEFKNFFYIISPG
HUM_GT1 WQCGGTLEIVTCSHV GHVFRKATPYTFPGG TGQIINKNNRRLAEV WMDEFKNFFYIISPG
COW_GT1 WQCGGTLEIVTCSHV GHVFRKATPYTFPGG TGQIINKNNRRLAEV WMDEFKNFFYIISPG
PIG_GT1 WQCGGTLEIVTCSHV GHVFRKATPYTFPGG TGQIINKNNRRLAEV WMDEFKTFFYIISPG
47169   WMCGGRMEDIPCSRV GHIYRKYVPYKVP-- AGVSLARNLKRVAEV WMDEYAEYIYQRRPE
        : ::: :: :.:::: ::::::. ::..:    :  : :: :::::: ::::: .:.:   :

MUS_GT1 VTKVDYGDISSRLGL RRKLQCKPFSWYLEN IYPDS-----QIPRH YFSLGEIRNVETNQC
RAT_GT1 VTKVDYGDISSRVGL RHKLQCKPFSWYLEN IYPDS-----QIPRH YFSLGEIRNVETNQC
HUM_GT1 VTKVDYGDISSRVGL RHKLQCKPFSWYLEN IYPDS-----QIPRH YFSLGEIRKEETNQC
COW_GT1 VTKVDYGDISSRLGL RHKLQCRPFSWYLEN IYPDS-----QIPRH YFSLGEIRNVETNQC
PIG_GT1 VTKVDYGDISSRLGL RHKLQCRPFSWYLEN IYPDS-----QIPRH YSSLGEIRNVETNQC
47169   YRHLSAGDVAVQKKL RSSLNCKSFKWFMTK IAWDLPKFYPPVEPP AAAWGEIRNVGTGLC
        : :: .::: :  : :..::::.:.::::.: :. :       :    .. .:::::: :  :

MUS_GT1 LDNMARKENEKVGIF NCHGMGG------NQ VFSYTANKEIR---- --TDDLCLDVSKLNG
RAT_GT1 LDNMARKENEKVGIF NCHGMGG------NQ VFSYTANKEIR---- --TDDLCLDVSKLNG
HUM_GT1 LDNMARKENEKVGIF NCHGMGG------NQ VFSYTANKEIR---- --TDDLCLDVSKLNG
COW_GT1 LDNMARKENEKVGIF NCHGMGG------NQ VFSYTANKEIR---- --TDDLCLDVSKLNG
PIG_GT1 LDNMARKENEKVGIF NCHGMGG------NQ VFSYTANKEIR---- --TDDLCLDVSKLNG
47169   ADTKHGALGSPLRLE GCVRGRGEAAWNNMQ VFTFTWREDIRPGDP QHTKKFCFDAISHTS
        ::. .    . : :   :. : :       : ::::::..:::       :...:.::  ..
```

Fig. 5C

```
MUS_GT1 PVTMLKCHHLKGNQL WEYDPVKLTLQHVNS NQCLDKATEEDSQVP SIRDCTGSRSQQWLL
RAT_GT1 PVTMLKCHHLKGNQL WEYDPVKLTLQHVNS NQCLDKATEEDSQVP SIRDCTGSRSQQWLL
HUM_GT1 PVTMLKCHHLKGNQL WEYDPVKLTLQHVNS NQCLDKATEEDSQVP SIRDCNGSRSQQWLL
COW_GT1 PVTMLKCHHLKGNQL WEYDPVKLTLQHVNS NQCLDKATDEDSQVP SIRDCSGSRSQQWLL
PIG_GT1 PVTMLKCHHLKGNQL WEYDPVKLTLQHVNS NQCLDKATEEDSQVP SIRDCSGSRSQQWLL
47169   PVTLYDCHSMKGNQL WKYR-KDKTLYHPVS GSCMDCS-ESDHRIF MNTCNPSSLTQQWLF
        :::::.::.:::::: :.:.  . ::.:  :  .::::.  :.:..:    .... : :::::.

MUS_GT1 R--NVTLPEIF---
RAT_GT1 R--NVTLPEIF---
HUM_GT1 R--NVTLPEIF---
COW_GT1 R--NVTLPEIF---
PIG_GT1 R--NVTLPEIF---
47169   EHTNSTVLEKFNRN
        .  : :: : :
```

Fig. 5D

```
gly-5b ----MIIFKKKAILK VLLLVPVFWICSLIF FAATSNDSSQIGSNN DLANK--IAEANFHP
gly-6a --MIASLIRSRRRSR RCVVYSVFLFGFLAL WGSFALALVFLSDMY IGEDQ--ISTQKAIK
gly-3  --MLSVGGGRSAVCR AVIATSIVWLLIDVV ILFYYLDPSTSQQQP FPEDNRILNRARRIE
gly-4  -----MLPRMLKMKT VGTVLAVIWLFGLAF IYVQSTSSSLRPPGR HPPPLP--QLDPLIP
gly-9  -----MLRYIIPRKK GTFVIAAFLTVAFFC IVAYHRNDRRRTKFQ FPDIEKYAEELVRLP
47169  ------MRRKEKRLL QAVALVLAALVLLPN VGLWALYRERQPDGT PGGSGAAVAPAAGQG

gly-5b KAAKQDVIQGFGPPI EPEPVVENNKVEEEE QPGGNLAKPKFMVDP NDPIYKKGDAAQAGE
gly-6a PIARSNYHVVVGHYN GNLPEDKKRNLTSEE LNAN-------LYAP HD---------DWGE
gly-3  PLPPAAQHDSDPDAH PIQPEKQEKQVYPVD KETANQLRKLMETQA FGPG-------YHGQ
gly-4  QNPPQNDEIRPKKSA PPIPTINLAEDTTIH ERTEKD----VT--- ------------WKT
gly-9  ETWNGELHQIPNYTA PR--EGP-------- ---------------- ------------GE
47169  SHSRQKKTFFLGDGQ KLKDWHDKEAIRRDA QRVGN---------- ------------GE

gly-5b LGKAVVVDKTKLSTE EKAKYDKGMLNNAFN QYASDMISVHRTLPT NIDAECKT---EKYN
gly-6a GG--AGVS--HLTPE QQKLADSTFAVNQFN LLVSDGISVRRSLPE IRKPSCRN---MTYP
gly-3  GGTGVTVP-----ED KKTIKEKRFLENQFN VVASEMISVNRTLPD YRSDACRTSGNNLKT
gly-4  FDVEKFLN-----KG KWHQGEDKYKANSFN QEASDALNPTRKIPD SREPQCRD--VDYSK
gly-9  KGKPVVLTG-----K DAELGQADMKKWFMN VHASDKISLDRDVPD PRIQACKD--IKYDY
47169  QGRPYPMT------- DAERVDQAYRENGFN IYVSDKISLNRSLPD IRHPNCNS---KRYL
                       .  .  : ::  . ::: ::: : :::  :   ::
```

Fig. 6A

Patent Application Publication


```
gly-5b ENLPRTSVIICFHNE AWSVLLRTVHSVLER TPDHLLEEVVLVDDF SDMDHTKRP-LEEYM
gly-6a DNLPTTSVIIVYHNE AYSTLLRTVWSVIDR SPKELLKEIILVDDF SDREFLRYPTLDTTL
gly-3  AGMPKTSIIIVFHNE AWTTLLRTLHSVINR SPRHLLEEIILVDDK SDRDYLVKP-LDSYI
gly-4  VGMQPTTVIITYHNE ARSSLLRTVFSVFNQ SPEELLLEIVLVDDN SQDVEIGKE-L----
gly-9  AALPKTSVIIIFTDE AWTPLLRTVHSVINR SPPELLQEVILLDDN SKRQELQEP-LDEHI
47169  ETLPNTSIIIPFHNE GWSSLLRTVHSVLNR SPPELVAEIVLVDDF SDREHLKKP-LEDY-
         ::.:::::.:::: :::.:::::::::::: :: .:: :::::::: :... :..: :: ::

gly-5b --SQFGGKVKILRME KREGLIRARLRGAAV ATGEVLTYLDSHCEC MEGWMEPLLDRIKRD
gly-6a --KPLPTDIKIIRSK ERVGLIRARMMGAQE AQGDVLTFLDSHCEC TKGWLEPLLTRIKLN
gly-3  --KMFPIPIHLVHLE NRSGLIRARLTGSEM AKGKILLFLDAHVEV TDGWLEPLVSRVAED
gly-4  ---AQIQRITVLRNN QREGLIRSRVKGAQV ARAPVLTFLDSHIEC NQKWLEPLLARIAEN
gly-9  --KRFGGKVRLIRKH DRHGLIRAKLAGARE AVGDIIVFLDSHCEA NHGWLEPIVQRISDE
47169  --MALFPSVRILRTK KREGLIRTRMLGASV ATGDVITFLDSHCEA NVNWLPPLLDRIARN
             ::::: . .: ::::.:: ::   : :.::.::::::.:.    ::.::: :: ..

gly-5b PTTVVCPVIDVIDDN TFEYHHSKAYFTSV- -GGFDWGLQFNWHSI P-ERDRKNRTRPIDP
gly-6a RKAVPCPVIDIINDN TFQYQKGIEMFR--- -GGFNWNLQFRWYGM PTAMAKQHLLDPTGP
gly-3  RKRVVAPIIDVISDD TFEYVTASETTWG-- --GFNWHLNFRWYAV PKRELNRRGSDRSMP
gly-4  PKAVVAPIIDVINVD NFNYVGASADLR--- -GGFDWTLVFRWEFM NEQLRKERHAHPTAP
gly-9  RTAIVCPMIDSISDN TLAYHGDWSLST--- -GGFSWALHFTWEGL S-EEEQKRRTKPTDY
47169  RKTIVCPMIDVIDHD DFRYETQAGDAMR-- -GAFDWEMYYKRIPI P---PELQKADPSDP
       .. ::.::::.:... .. :             :::.: : :.:  : .      :   .: .
```

Fig. 6B

```
gly-5b VRSPTMAGGLFSIDK EYFEKLGTYDPGFDI WGGENLELSFKIWMC GGTLEIVPCSHVGHV
gly-6a IESPTMAGGLFSINR NYFEELGEYDPGMDI WGGENLEMSFRIWQC GGRVEILPCSHVGHV
gly-3  IQTPTIAGGLFAIDK QFFYDIGSYDEGMQV WGGENLEISFRVWMC GGSLEIHPCSRVGHV
gly-4  IRSPTMAGGLFAISK EWFNELGTYDLDMEV WGGENLEMSFRVWQC GGSLEIMPCSRVGHV
gly-9  IRSPTMAGGLLAANR EYFFEVGGYDEEMDI WGGENLEISFRAWMC GGSIEFIPCSHVGHI
47169  FESPVMAGGLFAVDR KWFWELGGYDPGLEI WGGEQYEISFKVWMC GGRMEDIPCSRVGHI
       ..::.:::::..:.: .::.::: ::  ..: :::::.:::::::.: ::.:: :::::::::

gly-5b FRKRSPYKWR---TG VNVLKRNSIRLAEVW LDDYKTYYYERIN-- NQLGDFGDISSRKKL
gly-6a FRKSSPHDFPGK-SS GKVLNTNLLRVAEVW MDDWKHYFYKIAPQA HRMRSSIDVSERVEL
gly-3  FRKQTPYTFPG--GT AKVIHHNAARTAEVW MDEYKAFFYKMVP-A ARNVEAGDVSERKKL
gly-4  FRKKHPYTFPG--GS GNVFQKNTRRAAEVW MDEYKAIYLKNVP-S ARFVNFGDITDRLAI
gly-9  FRAGHPYNMTGRNNN KDVHGTNSKRLAEVW MDDYKRLYYMHRED- LRTKDVGDLTARHEL
47169  YRKYVPYKVP----A GVSLARNLKRVAEVW MDEYAEYIYQRRP-E YRHLSAGDVAVQKKL
       ::.  ::...       ... .:  :.:::: ::::. ....       :    :::. :..:

gly-5b REDLGCKSFKWYLDN IYP-ELFVPGESVAK GEVRNSAVQPARCLD CMVGRHEKNRPVGTY
gly-6a RKKLNCKSFKWYLQN VFQ-DHFLPTPLDRF GRMTSSS-NSSVCLA WTLRSSGIKTASTAD
gly-3  RETLQCKSFKWYLEN IYP-EAPLPADFRSL GAIVN----RFTEKC VDTNGKKDGQAPGIQ
gly-4  RDRLQCKSFKWYLEN VYP-QLEIPRKTPGK SFQMKIG-----NLC LDSMARKESEAPGLF
gly-9  RKRLNCKPFKWFLDN IAKGKFIMDEDVVAY GALHTVVSGTRMCTD TLQRDEKMSQLLGVF
47169  RSSLNCKSFKWFMTK IAWDLPKFYPPVEPP AAAWGEIRNVGTGLC ADTKHGALGSPLRLE
       : .:::::::::::.: :.     .                                    .
```

Fig. 6C

```
gly-5b QCHGQGGN------Q YWMLSKDGEIRRDE- -------SCVDYAGS DVMVFPCHGMKGNQE
gly-6a CLKIFHKT------Q LWLYTGDRRIRTDEH -LCLSVVQLLHTTSD WKIQLKECAGFDTEY
gly-3  ACHGAGGN------Q AWSLTGKGEIRSDDL -CLSSGHVYQIGSEL KLERCSVSKINVKHV
gly-4  GCHGTGGN------Q EWVFDQLTKTFKNAI -----SQLCLDFSSN TENKTVTMVKCENLR
gly-9  HCQGKGSSP-----Q LMSLSKEGNLRREN- --------TCASEEN GNIRMKTCSKKAQFN
47169  GCVRGRGEAAWNNMQ VFTFTWREDIRPGDP QHTKKFCFDAISHTS PVTLYDCHSMKGNQL
         .             : ..... . ... .       .     ....

gly-5b WRYNHDTGRLQHAVS QKCLG-------MTK DGAKLEMVACQYDDP YQHWKFKEYNEAKAI
gly-6a WDFKPKIGRFQNRKT GLCLASPDIFDPTKD EFNPPIVQKCRSSND RQNWTITEMSWLPEH
gly-3  FVFDDQAGTLLHKKT GKCVTG-------AD QRVTLDECGLGRKD- -QMWQLEGYQSP---
gly-4  PDTMVVEKNGWLTQG GKCLTVNQG----SG GDWLIYGAHCELNNG AQRWIFEKLDTYE--
gly-9  ERWAYENKMIRNLKS GKCMSTAN-----LK PGDNAIVVECDEKDE HQKWNFIDPAKA---
47169  WKYRKDK-TLYHPVS GSCMDCS------ES DHRIFMNTCNPSSLT -QQWLFEHTNSTVLE
         .             . ..::.                          :.: ..

gly-5b EHGAKPPS
gly-6a P-------
gly-3  --------
gly-4  --------
gly-9  --------
47169  KFNRN---
```

Fig. 6D